

On the Convergence of the
Iteration Sequence in Primal-Dual
Interior-Point Methods

R.A. Tapia
Y. Zhang
Y. Ye

August, 1991
(Revised August 1993)

TR91-24

Report Documentation Page				Form Approved OMB No. 0704-0188	
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE AUG 1993		2. REPORT TYPE		3. DATES COVERED 00-00-1993 to 00-00-1993	
4. TITLE AND SUBTITLE On the Convergence of the Iteration Sequence in Primal-Dual Interior-Point Methods				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Computational and Applied Mathematics Department ,Rice University,6100 Main Street MS 134,Houston,TX,77005-1892				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 26	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

On the Convergence of the Iteration Sequence in Primal-Dual Interior-Point Methods *

Richard A. Tapia [†], Yin Zhang [‡] and Yinyu Ye [§]

August, 1991

(Revised August 1993)

*This research was supported in part by NSF Coop. Agr. No. CCR-8809615. A part of this research was performed in June, 1991 while the second and the third authors were at Rice University as visiting members of the Center for Research in Parallel Computation.

[†]Department of Mathematical Sciences and the Center for Research in Parallel Computation, Rice University, Houston, Texas 77251-1892. Research supported in part by AFOSR 89-0363, DOE DEFG05-86ER25017 and ARO 9DAAL03-90-G-0093.

[‡]Department of Mathematics and Statistics, University of Maryland, Baltimore County Campus, Baltimore, Maryland 21228. Research supported in part by NSF DMS-9102761 and DOE DE-FG05-91ER25100.

[§]Department of Management Sciences, The University of Iowa, Iowa City, Iowa 52242. Research supported in part by NSF DDM-8922636.

Abstract

Recently, numerous research efforts, most of them concerned with superlinear convergence of the duality gap sequence to zero in the Kojima-Mizuno-Yoshise primal-dual interior-point method for linear programming, have as a primary assumption the convergence of the iteration sequence. Yet, except for the case of nondegeneracy (uniqueness of solution), the convergence of the iteration sequence has been an important open question now for some time. In this work we demonstrate that for general problems, under slightly stronger assumptions than those needed for superlinear convergence of the duality gap sequence (except of course the assumption that the iteration sequence converges), the iteration sequence converges. Hence, we have not only established convergence of the iteration sequence for an important class of problems, but have demonstrated that the assumption that the iteration sequence converges is redundant in many of the above mentioned works.

Keywords: Linear programming, Primal-dual interior-point algorithms, Convergence of iteration sequence.

Abbreviated Title: Convergence of Interior-point Methods

1 Introduction

This paper considers linear programs in the standard form:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned} \tag{1.1}$$

where $c, x \in \mathbf{R}^n$, $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n}$ ($m < n$) and A has full rank m .

The first-order optimality conditions for (1.1) can be written as the following $2n$ by $2n$ nonlinear system with non-negativity constraints on the variables:

$$F(x, y) = \begin{pmatrix} Ax - b \\ By - Bc \\ XYe \end{pmatrix} = 0, \quad (x, y) \geq 0, \tag{1.2}$$

where $B \in \mathbf{R}^{(n-m) \times n}$ is any matrix such that the columns of B^T form a basis for the null space of A , $X = \text{diag}(x)$, $Y = \text{diag}(y)$ and e is the n -vector of all ones.

The feasibility set of problem (1.2) is defined as

$$\mathcal{F} = \{(x, y) : x, y \in \mathbf{R}^n, Ax = b, By = Bc, (x, y) \geq 0\}.$$

A feasible pair $(x, y) \in \mathcal{F}$ is said to be strictly feasible if it is positive. In this work we assume that strictly feasible points exist. We denote the solution set of problem (1.2) by

$$\mathcal{S} = \{(x^*, y^*) : F(x^*, y^*) = 0, (x^*, y^*) \geq 0\}.$$

The algorithms under consideration in this study are primal-dual interior-point algorithms as originally proposed by Kojima, Mizuno and Yoshise in [10] and are described by the following algorithmic framework.

Algorithm 1 (Primal-Dual Algorithm)

Given a strictly feasible pair (x^0, y^0) . For $k = 0, 1, 2, \dots$, do

(1) *Choose $\sigma^k \in [0, 1)$ and set $\mu^k = \sigma^k \frac{(x^k)^T y^k}{n}$.*

(2) Solve the following system for $(\Delta x^k, \Delta y^k)$:

$$F'(x^k, y^k) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -F(x^k, y^k) + \mu^k \begin{pmatrix} 0 \\ e \end{pmatrix}.$$

(3) Choose $\tau^k \in (0, 1)$ and set $\alpha^k = \min(1, \tau^k \hat{\alpha}^k)$ where

$$\hat{\alpha}^k = \frac{-1}{\min((X^k)^{-1} \Delta x^k, (Y^k)^{-1} \Delta y^k)}. \quad (1.3)$$

(4) Form the new iterate

$$(x^{k+1}, y^{k+1}) = (x^k, y^k) + \alpha^k (\Delta x^k, \Delta y^k).$$

Note that the choice of step-length α^k guarantees $(x^{k+1}, y^{k+1}) > 0$. Moreover, it can be easily verified that

$$(x^{k+1})^T y^{k+1} = (1 - \alpha^k (1 - \sigma^k)) (x^k)^T y^k < (x^k)^T y^k. \quad (1.4)$$

Immediately following the presentation of the primal-dual algorithm, one group of researchers was mainly concerned with choices of the parameters in Algorithm 1 that led to good complexity bounds, while another group was concerned with choices of these parameters that led to numerically effective algorithms. Practical implementations and numerical experiments performed by MacShane, Monma, and Shanno [13] and by Lustig, Marsten, and Shanno [11], for example, demonstrated that primal-dual algorithms can be computationally very effective.

The fast convergence of the latter algorithms, in part, motivated the recent flurry of papers concerned with establishing superlinear convergence of the duality gap to zero. For linear programming applications, these papers include Zhang, Tapia, and Dennis [24] and Zhang and Tapia [21]. In addition, Ye, Tapia, and Zhang [20], McShane [12], Ye, Güler, Tapia, and Zhang [19] and Mehrotra [15] recently studied the Mizuno-Todd-Ye predictor-corrector algorithm that takes $\sigma^k = 1$ and $\sigma^k = 0$ alternatively. For linear complementarity problems, these papers include Zhang, Tapia, and Potra [25], Kojima, Kurita, and Mizuno [9]

and Ji, Potra, Tapia, and Zhang [8]. In the context of the present work, it is important to observe that in establishing superlinear convergence of the duality gap (or complementarity) to zero, all of the above authors, except Ye *et al* [19], and Mehrotra [15] assumed the convergence of the iteration sequence. For the Mizuno-Todd-Ye predictor-corrector algorithm, Ye *et al* [19] and independently Mehrotra [15] obtained the impressive result of quadratic convergence, when counting the predictor and corrector steps as a single step, without the assumption of nondegeneracy or the assumption of convergence of the iteration sequence. Whether such convergence can be obtained for Algorithm 1 without these assumptions is an important question and is the subject of current research. It is also of concern to us that while in practice the convergence of the iteration sequence is always observed, there is, to our knowledge, no theory in the literature implying that this should be expected for problems that do not have unique solutions. It is the objective of the present paper to partially fill this conspicuous gap.

This paper is organized as follows. Section 2 contains notation and some known results that will be used in the development of this paper. In Section 3, we first establish several technical results. These technical results then are used to derive a bound on the distance from the $(k+1)$ -th iterate to a solution point in terms of the distance from the k -th iterate to this solution point and the centering parameter σ_k . The parametric form of this inequality bears a striking similarity to the form of the bounded deterioration principle for secant updates used by Broyden, Dennis and Moré [1] to establish convergence for several secant methods. For this reason, we shall refer to the distance inequality established in Theorem 3.1 as the bounded deterioration of the iteration sequence. We then derive Theorem 3.2, a useful bound on the search steps of Algorithm 1, from the bounded deterioration of the iteration sequence and the Hoffman lemma [7]. We present the main theorem of this paper, Theorem 4.1, in Section 4. It states that, under mild assumptions, the iteration sequence generated by Algorithm 1 converges. Finally, we make some concluding remarks in Section 5.

2 Preliminaries

In this section, we introduce notation and some known results. For brevity, we use the notation $z = (x, y) \in \mathbf{R}^{2n}$, i.e.,

$$z_i = \begin{cases} x_i, & 1 \leq i \leq n, \\ y_i, & n < i \leq 2n. \end{cases} \quad (2.1)$$

Let $\hat{e} = (0 \dots 0 \ 1 \dots 1)^T \in \mathbf{R}^{2n}$ where the numbers of zeros and ones are both n . Then Algorithm 1 can be written as the following iterative process:

$$z^{k+1} = z^k - \alpha^k [F'(z^k)]^{-1} (F(z^k) - \mu^k \hat{e}). \quad (2.2)$$

A straightforward calculation gives

$$F'(x, y) = \begin{bmatrix} A & 0 \\ 0 & B \\ Y & X \end{bmatrix}. \quad (2.3)$$

The following proposition is known; however we could not locate a simple proof for the case of linear programming in the literature and therefore, for the sake of completeness, we include a short proof. Observe that no assumptions are made on the solution set \mathcal{S} of Problem (1.2). Recall that by the distance from a point z to a set \mathcal{S} , we mean

$$\text{dist}(z, \mathcal{S}) = \inf \{ \|z - s\| : s \in \mathcal{S} \}.$$

Proposition 2.1 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1. Then*

- (i) $\{z^k\}$ is bounded.
- (ii) If $(x^k)^T y^k \rightarrow 0$, then $\text{dist}(z^k, \mathcal{S}) \rightarrow 0$.

Proof: (i) Since $A(x^k - x^0) = 0$, $B(y^k - y^0) = 0$ and the columns of B^T form a basis for the null space of A ,

$$0 = (x^k - x^0)^T (y^k - y^0) = (x^k)^T y^k + (x^0)^T y^0 - (y^0)^T x^k - (x^0)^T y^k.$$

From (1.4),

$$(y^0)^T x^k + (x^0)^T y^k = (x^k)^T y^k + (x^0)^T y^0 \leq 2(x^0)^T y^0.$$

The boundedness of (x^k, y^k) now follows from this inequality and the fact that both (x^0, y^0) and (x^k, y^k) are positive.

(ii) Let \mathcal{L} denote the set of limit points of $\{z^k\}$. Since the duality gap converges to zero, $\mathcal{L} \subset \mathcal{S}$. Given any $\epsilon > 0$, cover \mathcal{L} with

$$\mathcal{B}(\epsilon) = \bigcup_{z^* \in \mathcal{L}} \{z : \|z - z^*\| < \epsilon\}.$$

Since $\{z^k\}$ is bounded, all but a finite number of z^k 's are in $\mathcal{B}(\epsilon)$; otherwise, $\{z^k\}$ would have a new limit point that is not in \mathcal{L} , and this would be a contradiction. Hence, there exists a smallest number N such that $z^k \in \mathcal{B}(\epsilon)$ for all $k \geq N$. This implies that $\text{dist}(z^k, \mathcal{S}) < \epsilon$ for all $k \geq N$ and completes the proof. \square

Lemma 2.1 below is Lemma 2 of Güler and Ye [5], tailored to fit the needs of this paper. Many interesting properties of interior-point algorithms follow from this simple fact.

Lemma 2.1 (Güler-Ye)

Let $\{z^k\} = \{(x^k, y^k)\}$ be generated by Algorithm 1. Assume

A1 $(x^k)^T y^k$ converges to zero.

A2 $\min(X^k Y^k e) / (x^k)^T y^k \geq \gamma/n$ for all k and some $\gamma \in (0, 1)$.

Let z^ be a limit point of $\{z^k\}$. Then*

(i) z^* is a strictly complementary solution of Problem 1.2 and

(ii) $\lim_{k \rightarrow \infty} \inf z_i^k > 0$ for every i such that $z_i^* > 0$.

Recently, El-Bakry, Tapia, and Zhang [2] demonstrated that the relative interior of \mathcal{S} and the set of solutions satisfying strict complementarity coincide. Furthermore, the zero-nonzero structure of solutions in the relative interior of \mathcal{S} (equivalently solutions satisfying strict complementarity) is invariant. Hence, under Assumption A2 of Lemma 2.1, the limit points of the iteration sequence generated by Algorithm 1 have considerable structure.

For any $z^* = (x^*, y^*)$ in the relative interior of \mathcal{S} , define

$$I_x^+ = \{i : x_i^* > 0, 1 \leq i \leq n\} \quad \text{and} \quad I_y^+ = \{i : y_i^* > 0, 1 \leq i \leq n\}. \quad (2.4)$$

Since the zero-nonzero structure of the relative interior of \mathcal{S} is invariant, the above two index sets are independent of the choice of z^* . By strict complementarity of z^* ,

$$I_x^+ \cup I_y^+ = \{1, 2, \dots, n\} \quad \text{and} \quad I_x^+ \cap I_y^+ = \emptyset. \quad (2.5)$$

3 Bounded Deterioration

We consider the following three matrices which were introduced in Zhang, Tapia, and Dennis [24]:

$$P = I - X^{\frac{1}{2}}Y^{-\frac{1}{2}}A^T(AXY^{-1}A^T)^{-1}AX^{\frac{1}{2}}Y^{-\frac{1}{2}}, \quad (3.1)$$

$$H_p = (XY)^{-\frac{1}{2}}P(XY)^{-\frac{1}{2}}, \quad (3.2)$$

$$H_d = (XY)^{-\frac{1}{2}}(I - P)(XY)^{-\frac{1}{2}}. \quad (3.3)$$

It is worth noting that P and $I - P$ are orthogonal projection matrices.

Lemma 3.1 *Let H_p and H_d be defined by (3.2) and (3.3), respectively, then*

$$F'(x, y)^{-1} \begin{pmatrix} 0 \\ w \end{pmatrix} = \begin{pmatrix} XH_p w \\ YH_d w \end{pmatrix}.$$

Proof: The proof follows from Proposition 2.1 in Zhang, Tapia, and Dennis [24]. \square

Lemma 3.2 *Let $(x^*, y^*) \in \mathcal{S}$ and $(x, y) \in \mathcal{F}$. Then*

$$F'(x, y) \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} - F(x, y) = \begin{pmatrix} 0 \\ (X - X^*)(Y - Y^*)e \end{pmatrix}.$$

Proof: The proof follows from direct substitution. \square

Lemma 3.3 Let $z^* = (x^*, y^*) \in \mathcal{S}$, $z = (x, y) \in \mathcal{F}$ and $z^+ = (x^+, y^+)$ be given by

$$z^+ = z - \alpha[F'(z)]^{-1}(F(z) - \mu\hat{e}).$$

Then

$$z^+ - z^* = \begin{pmatrix} XH_p w \\ YH_d w \end{pmatrix},$$

where

$$w = (X - X^*)(Y - Y^*)e + (1 - \alpha)XYe + \alpha\mu\hat{e}. \quad (3.4)$$

Proof: Using Lemma 3.2, we have

$$\begin{aligned} z^+ - z^* &= z - z^* - \alpha[F'(z)]^{-1}(F(z) - \mu\hat{e}) \\ &= [F'(z)]^{-1}[F'(z)(z - z^*) - F(z) + (1 - \alpha)F(z) + \alpha\mu\hat{e}] \\ &= [F'(z)]^{-1} \begin{pmatrix} 0 \\ w \end{pmatrix}. \end{aligned}$$

Now the lemma follows from Lemma 3.1. \square

Lemma 3.4 Let P_i be the i -th row (or column) of an orthogonal projection matrix P . Then

$$\|P_i\|_2 \leq 1.$$

Proof: Since $I - P$ is symmetric and positive semi-definite, $P_{ii} \leq 1$. Now $P^2 = P$ leads to $\|P_i\|_2^2 = P_{ii} \leq 1$. \square

We now derive our bounded deterioration relationship between the distances from two consecutive iterates, z^k and z^{k+1} , to any solution z^* .

Theorem 3.1 Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1 and let $z^* = (x^*, y^*)$ be any solution to Problem 1.2. Assume

A1 $(x^k)^T y^k \rightarrow 0$.

A2 $\min(X^k Y^k e) / (x^k)^T y^k \geq \gamma/n$ for all k and some $\gamma \in (0, 1)$.

Then there exist constants $\beta'_i > 0$ and $\beta''_i > 0$ such that

$$|z_i^{k+1} - z_i^*| \leq (1 + z_i^k \beta'_i) \|z^k - z^*\| + z_i^k \beta''_i \sigma^k, \quad i = 1, 2, \dots, 2n. \quad (3.5)$$

Proof: For simplicity, we will replace the superscript $(k+1)$ by $+$ and drop the superscript k . We let $\hat{z} = z + \Delta z$. Since $z^+ = z + \alpha \Delta z$ and $\alpha \in (0, 1]$, for every index i we have

$$z_i^+ \in [\min(z_i, \hat{z}_i), \max(z_i, \hat{z}_i)].$$

Therefore, for z^*

$$|z_i^+ - z_i^*| \leq \max(|z_i - z_i^*|, |\hat{z}_i - z_i^*|). \quad (3.6)$$

Without loss of generality, let us consider $i \leq n$, i.e., $z_i = x_i$, $z_i^* = x_i^*$ and so on. The proof for $i > n$ is the same.

Letting $\alpha = 1$ in Lemma 3.3 and using the definition of H_p and w , we obtain

$$\hat{x}_i - x_i^* = x_i \sum_{j=1}^n (x_i y_i)^{-\frac{1}{2}} P_{ij}(x_j y_j)^{-\frac{1}{2}} w_j = x_i (S'_i + S''_i), \quad (3.7)$$

where

$$S'_i = \sum_{j=1}^n (x_i y_i)^{-\frac{1}{2}} P_{ij}(x_j y_j)^{-\frac{1}{2}} (x_j - x_j^*)(y_j - y_j^*),$$

and

$$S''_i = \sigma \sum_{j=1}^n (x_i y_i)^{-\frac{1}{2}} P_{ij}(x_j y_j)^{-\frac{1}{2}} \frac{x_j^T y_j}{n}.$$

In S''_i we used the relation $\mu = \sigma x^T y / n$. Let us consider the above two sums. Notice that

$$(x_j - x_j^*)(y_j - y_j^*) = \begin{cases} y_j(x_j - x_j^*), & \text{if } j \in I_x^+, \\ x_j(y_j - y_j^*), & \text{if } j \in I_y^+. \end{cases}$$

We lose no generality by assuming that (x^*, y^*) satisfies strict complementarity. For if (3.5) holds for all z^* in the relative interior of the solution set S , then by continuity of the absolute value and norm, it holds for all points on the relative boundary of S . As previously mentioned at the end of Section 2, the points in the relative interior of S are exactly the

solutions satisfying strict complementarity. By strict complementarity, for each index j either $x_j^* > 0$ or $y_j^* > 0$. Therefore,

$$S'_i = \left[\sum_{j \in I_x^+} P_{ij} \left(\frac{x_j y_j}{x_i y_i} \right)^{\frac{1}{2}} \frac{x_j - x_j^*}{x_j} + \sum_{j \in I_y^+} P_{ij} \left(\frac{x_j y_j}{x_i y_i} \right)^{\frac{1}{2}} \frac{y_j - y_j^*}{y_j} \right].$$

Notice that

$$\frac{x_j y_j}{x_i y_i} = \frac{x_j y_j}{x^T y} \frac{x^T y}{x_i y_i} \leq \frac{n}{\gamma}$$

is bounded above and $\{x_j, j \in I_x^+\}$ and $\{y_j, j \in I_y^+\}$ are bounded away from zero (see (ii) of Lemma 2.1). Moreover, the matrix P is an orthogonal projection matrix; by Lemma 3.4, its rows are bounded (independent of k). Therefore, we can find $\beta'_i > 0$ such that

$$|S'_i| \leq \beta'_i \|z - z^*\|.$$

In addition, by Assumption A2 and Lemma 3.4 again, we can find $\beta''_i > 0$ such that

$$|S''_i| = \sigma \left(\frac{x^T y / n}{x_i y_i} \right)^{\frac{1}{2}} \left| \sum_{j=1}^n P_{ij} \left(\frac{x^T y / n}{x_j y_j} \right)^{\frac{1}{2}} \right| \leq \beta''_i \sigma.$$

It follows from (3.7) that

$$|\hat{x}_i - x_i^*| \leq x_i (|S'_i| + |S''_i|) \leq x_i (\beta'_i \|z - z^*\| + \beta''_i \sigma).$$

Finally, (3.5) follows from (3.6) and the above inequality. This completes the proof. \square

The following inequality, a consequence of the bounded deterioration and the Hoffman Lemma [7], will be a critical ingredient in the establishment of our convergence theory.

Theorem 3.2 *Let $\{z^k = (x^k, y^k)\}$ and $\{\Delta z^k\}$ be generated by Algorithm 1. Assume*

A1 $(x^k)^T y^k \rightarrow 0$.

A2 $\min(X^k Y^k e) / (x^k)^T y^k \geq \gamma / n$ for all k and some $\gamma \in (0, 1)$.

Then there exist constants $\beta' > 0$ and $\beta'' > 0$ such that

$$\|\Delta z^k\| \leq \beta' (x^k)^T y^k + \beta'' \sigma^k. \quad (3.8)$$

and, in particular,

$$\|z^{k+1} - z^k\| \leq \beta' (x^k)^T y^k + \beta'' \sigma^k. \quad (3.9)$$

Proof: Since $z^{k+1} = z^k + \alpha^k \Delta z^k$ and $\alpha^k \in (0, 1]$, it suffices to prove (3.8). Notice that the bounded deterioration (3.5) holds for any $\alpha^k \in (0, 1]$. When $\alpha^k = 1$, from bounded deterioration (3.5) we have for any $z^* \in \mathcal{S}$

$$\|\Delta z^k\| \leq \|(z^k + \Delta z^k) - z^*\| + \|z^k - z^*\| \leq \hat{\beta}\|z^k - z^*\| + \beta''\sigma^k. \quad (3.10)$$

Now, consider the system of linear inequalities

$$\begin{aligned} Ax - b &\leq 0 \\ -Ax - b &\leq 0 \\ By - Bc &\leq 0 \\ -By + Bc &\leq 0 \\ c^T x - b^T (AA^T)^{-1} A(c - y) &\leq 0 \\ -x &\leq 0 \\ -y &\leq 0. \end{aligned} \quad (3.11)$$

The system (3.11) is equivalent to the first-order conditions (1.2). To see this consider (x, y) satisfying (3.11). We have $B(y - c) = 0$. Since both A and B are full rank and $A^T = 0$, the columns of A^T form a basis for the null space of B . Hence there exists λ such that $y - c = A^T \lambda$. Therefore

$$\begin{aligned} c^T x - b^T (AA^T)^{-1} A(c - y) &= c^T x + b^T \lambda \\ &= c^T x + x^T A^T \lambda \\ &= x^T y. \end{aligned} \quad (3.12)$$

The second line above resulted from the substitution $b = Ax$. We have shown that the lengthy expression in (3.11) is merely a linear inequality requiring the duality gap $x^T y$ to be nonpositive. We know that for feasible points the duality gap is always nonnegative. Hence, this linear inequality is equivalent to $x^T y = 0$. We are now in a position to use the Hoffman Lemma [7] with the ℓ_2 norm for the choice of both positive homogeneous functionals. This lemma now says there exist a constant $\beta > 0$ and $w^k \in \mathcal{S}$ for every k such that

$$\|z^k - w^k\| \leq \beta(x^k)^T y^k.$$

Since $z^* \in \mathcal{S}$ is arbitrary in (3.10), it follows that (3.8) holds with $\beta' = \beta\hat{\beta}$. This proves the theorem. \square

Observe that when $\sigma^k = 0$, Theorem 3.2 reduces to Lemma 3.2 and Theorem 3.1 of Ye *et al* [19], i.e., a bound on the Newton step. Theorem 3.2 can also be proved using this bound on the Newton step and the bound on $|S_i''|$ in the proof of Theorem 3.1, or in perhaps other ways. We have chosen to make Theorem 3.2 a consequence of Theorem 3.1 because we believe that Theorem 3.1 is of interest in its own right.

4 Convergence Results

Theorem 4.1 below represents our main contribution to the convergence theory for the iteration sequence generated by primal-dual interior-point methods for linear programming.

Given any z^* in the relative interior of \mathcal{S} , let

$$I_0 = \{i : z_i^* = 0, \quad 1 \leq i \leq 2n\}.$$

From the remarks above, I_0 is invariant with respect to the choices of z^* .

Recall that a set in \mathbf{R}^n is said to be a continuum if it is closed and cannot be written as the disjoint union of two proper closed subsets. In \mathbf{R}^1 the continua are closed intervals. Clearly, a continuum is uncountable.

The following conditions will be selectively used in Theorem 4.1. Their compatibility will be demonstrated in Proposition 4.1.

C1 $(x^k)^T y^k \rightarrow 0;$

C2 $\min(X^k Y^k e) / (x^k)^T y^k \geq \gamma/n$ for all k and some $\gamma \in (0, 1)$.

C3 $\sigma^k \rightarrow 0.$

C4 $\sigma^k \rightarrow 0$ at least R -linearly.

C5 $\sigma^k = O((x^k)^T y^k).$

C6 $\tau^k \geq \tau$ for some $\tau \in (0, 1)$.

Theorem 4.1 *Let $\{z^k = (x^k, y^k)\}$ be generated by Algorithm 1. Then*

(i) *Conditions C1-C2 imply, $z_i^k \rightarrow 0$ for each $i \in I_0$.*

(ii) *Conditions C1-C3 imply,*

- *$\{z^k\}$ converges to some z^* in the relative interior of \mathcal{S} , or*
- *the limit points of $\{z^k\}$ form a continuum.*

(iii) *Conditions C1, C2, C4 and C6 imply, $\{z^k\}$ converges to some z^* in the relative interior of \mathcal{S} .*

(iv) *Conditions C1, C2, C5 and C6 imply, any Q or R convergence behavior present in the convergence of $\{(x^k)^T y^k\}$ to zero is reflected as R -convergence behavior in the convergence of $\{z^k\}$ to z^* .*

Proof: (i) Consider the component sequence $\{z_i^k\}$ for $i \in I_0$. It is bounded from (i) of Proposition 2.1. By Lemma 2.1 and El-Bakry, Tapia, and Zhang [2], all limit points of $\{z^k\}$ are in the relative interior of \mathcal{S} and have identical zero-nonzero structure; thus all limit points of $\{z_i^k\}$ are zero. This leads to $z_i^k \rightarrow 0$ for all $i \in I_0$.

(ii) Again let \mathcal{L} denote the set of limit points of $\{z^k\}$. Clearly, \mathcal{L} is compact. It is obviously closed and it is bounded since $\{z^k\}$ is bounded by (i) of Proposition 2.1. If \mathcal{L} is the singleton set $\{z^*\}$, then $\{z^k\}$ converges to z^* since it is a bounded sequence with only one limit point.

Now suppose that \mathcal{L} contains more than one point and is not a continuum. Then we can write $\mathcal{L} = \mathcal{L}_1 \cup \mathcal{L}_2$ where \mathcal{L}_i , $i = 1, 2$, is nonempty and compact. By Hausdorff separation we can find sets \mathcal{O}_i , $i = 1, 2$, such that \mathcal{O}_i is open, $\mathcal{L}_i \subset \mathcal{O}_i$ and the closures of \mathcal{O}_1 and \mathcal{O}_2 are disjoint. Clearly, \mathcal{O}_i contains an infinite number of z^k 's, or else, \mathcal{L}_i would be empty; also only a finite number of z^k 's are not in $\mathcal{O}_1 \cup \mathcal{O}_2$. This latter statement follows from the observation that if an infinite number of z^k 's are not contained in $\mathcal{O}_1 \cup \mathcal{O}_2$, then they would have a limit

point not in $\mathcal{O}_1 \cup \mathcal{O}_2$; and therefore not in \mathcal{L} . This would be a contradiction. It follows that we can select a subsequence $\{z^{k_j}\}$ with the property that $z^{k_j} \in \mathcal{O}_1$ and $z^{k_j+1} \in \mathcal{O}_2$. We can choose a thinner subsequence, which we also call $\{z^{k_j}\}$, with the property that $z^{k_j} \rightarrow z^*$ for some $z^* \in \mathcal{L}_1$ and $z^{k_j+1} \in \mathcal{O}_2$ for all j . However, this contradicts inequality (3.8) of Theorem 3.2.

(iii) It is known (see Zhang, Tapia, and Dennis [24], for example) that C2 and C6 imply $\alpha^k \geq \alpha > 0$ for all k and some α . From (1.4) it follows that the duality gap sequence converges Q -linearly to zero. It now follows from (3.8) and C4 that there exist positive constants η and $\delta < 1$ and a positive integer K such that

$$\|z^{k+1} - z^k\| \leq \eta\delta^k, \quad k \geq K. \quad (4.1)$$

In a standard manner, (4.1) implies that for all positive integers N

$$\|z^{k+N} - z^k\| \leq \frac{\eta}{1-\delta}\delta^k, \quad k \geq K. \quad (4.2)$$

Hence, $\{z^k\}$ is a Cauchy sequence and is therefore convergent.

(iv) The proof of (iv) follows from the derivation of the appropriate inequality analogous to (4.2) and letting $N \rightarrow \infty$. This proves the theorem. \square

The quantities τ^k and σ^k are under direct control of the algorithm designer. This is an extremely important point and often missed by some readers. Even without the knowledge of (iii), it seems that it would be quite pathological for one to choose $\sigma^k \rightarrow 0$ in any manner that did not reflect R -linear convergence.

It is worth pointing out the implications of Theorem 4.1 on the superlinear and quadratic convergence theory of Zhang, Tapia, and Dennis [24]. First, (iii) implies that the nondegeneracy assumption for quadratic convergence is no longer necessary. Second, (ii) implies that the assumption of the convergence of the iteration sequence made for superlinear convergence of the duality gap to zero is no longer necessary, due to the fact that in order to establish superlinear convergence one only needs $\|z^{k+1} - z^k\| \rightarrow 0$. These facts have been documented in Zhang and Tapia [23].

It is perhaps interesting to observe that in the proof of (ii) by partitioning \mathcal{L} so that \mathcal{L}_1 is a singleton set we have established that either $\{z^k\}$ converges or \mathcal{L} has no isolated points. In analysis, a closed set in which every point is a limit point of the set is called a perfect set. Clearly, \mathcal{L} is a perfect set in the case of nonconvergence. It is known that in a separable metric space, every perfect set must be uncountable. See Gelbaum and Olmsted [3] for a discussion and see Hausdorff [6] for a proof and further discussion. The interval $[0, 1]$ is a perfect set in \mathbf{R}^1 . One might conjecture that a perfect set in \mathbf{R}^1 contains a continuum, or at least has positive Lebesgue measure. However, in \mathbf{R}^1 the celebrated Cantor set is a perfect set of Lebesgue measure zero; hence it contains no continuum. It follows that this line of reasoning would produce a weaker result in (ii).

It is also interesting to note that (ii) of Theorem 4.1 has the flavor of a result established by Ostrowski [17] in the context of the gradient method.

We now demonstrate the compatibility of Conditions C1-C6 by showing that choices for σ^k and α^k exist such that these assumptions are simultaneously satisfied. This means that the algorithm designer can design an algorithm satisfying C1-C6 or any subset of these conditions.

Proposition 4.1 *For Algorithm 1 there exist parameter sequences $\{\sigma^k\}$ and $\{\alpha^k\}$, and the corresponding iteration sequence $\{(x^k, y^k)\}$ such that Conditions C1-C6 are satisfied.*

Proof: It suffices to demonstrate that Conditions C1, C2, C5 and C6 are simultaneously satisfied. Let

$$\sigma^k = \min(\sigma, \rho(x^k)^T y^k) \quad (4.3)$$

where $\sigma \in (0, 1)$ and $\rho > 0$. For each k choose τ^k so that $\alpha^k = \tau^k \hat{\alpha}^k$ is the largest number in $(0, 1]$ such that

$$\gamma \leq \frac{x_i^{k+1} y_i^{k+1}}{(x^{k+1})^T y^{k+1} / n} \leq \Gamma, \quad i = 1, 2, \dots, n, \quad (4.4)$$

for some fixed constants $\gamma \in (0, 1/2)$ and $\Gamma \in (2, n)$. It has been shown (see Lemma 3.4 of Zhang and Tapia [21]) that

$$\alpha^k \geq \min \left(1, \frac{\sigma^k (x^k)^T y^k / n}{2 \max |\Delta x_i^k \Delta y_i^k|} \right). \quad (4.5)$$

The proposition will be proved if we can show that $\{\alpha^k\}$ is bounded away from zero.

When $\sigma^k = \sigma$, it is known that $\alpha^k = \Omega(\frac{1}{n})$ (see [21], for example). Otherwise, $\sigma^k = \rho(x^k)^T y^k$. By Theorem 3.2, $|\Delta x_i^k \Delta y_i^k| = O((x^k)^T y^k)^2$. Hence it follows from (4.5) that α^k is also bounded away from zero. This completes the proof. \square

5 Concluding Remarks

The main result of this paper is Theorem 4.1. Among other things, it says that $\sigma^k \rightarrow 0$ will guarantee that the step $z^{k+1} - z^k$ converges to zero; but this alone does not imply the convergence of the iteration sequence. However, this convergence is guaranteed if the step converges to zero fast enough, i.e., at least R -linearly, and this in turn is guaranteed if σ^k converges to zero R -linearly.

We now argue the mildness of the assumptions made in Theorem 4.1. Condition C1, the convergence of the duality gap sequence to zero, is necessary if we are to consider convergence to a solution. Condition C2 is a crucial ingredient for polynomial convergence of the duality gap sequence to zero. Moreover, Conditions C2, C3 and $\tau^k \rightarrow 1$ are fundamental in establishing Q -superlinear convergence (see Zhang, Tapia, and Dennis [24]) of the duality gap sequence to zero.

We now summarize the situation as it concerns the convergence of the iteration sequence in primal-dual interior-point methods. If the problem is both primal and dual nondegenerate, then the solution is unique and the convergence of the iteration sequence readily follows. Numerous papers in the literature have assumed the convergence of the iteration sequence for degenerate problems in order to facilitate their particular application. Yet, prior to the present work there was no theory whatsoever addressing this issue. The present work shows, under mild assumptions, that if the centering parameter sequence is chosen so that $\{\sigma_k\}$ converges to zero at least R -linearly, then convergence of the iteration sequence results. This result is satisfying because it implies that the conditions guaranteeing superlinear convergence of the duality gap sequence to zero essentially imply convergence of the iteration

sequence; hence the assumption of iteration sequence convergence is essentially redundant in this theory. However, at this juncture we do not know what points in the solution set are acceptable as the limit points of convergent iteration sequences. It is our conjecture that the iteration sequence, under the conditions of our theory, does not necessarily converge to the analytic center of the solution set. Recently, while this paper was in review, Zhang and Tapia [22] demonstrated, under mild assumptions, that if $\{\sigma_k\}$ is bounded away from zero and the iteration sequence converges, then the convergence is necessarily to the analytic center. The present theory does not speak to this case. Hence, the convergence or lack of convergence of the iteration sequence for the case $\{\sigma_k\}$ is bounded away from zero is presently an annoying gap in our understanding. This important question was highlighted further when even more recently, also while the current paper was in review, Gonzaga and Tapia [4] showed that the iteration sequence generated by the Mizuno-Todd-Ye predictor-corrector primal-dual interior-point method converges, and the convergence is to the analytic center of the solution set. The Mizuno-Todd-Ye predictor-corrector method can be viewed as a special case of the Kojima-Mizuno-Yoshise method (Algorithm 1) where the parameter sequence $\{\sigma_k\}$ is chosen so that its terms alternate between the values 0 and 1. Hence the algorithm alternates between fidelity to the central path and fidelity to the pure Newton method. The Gonzaga-Tapia and the Zhang-Tapia theories do not take away from the present theory; they do, however, add more pieces to a very interesting and challenging puzzle.

We end this section, and the paper, by briefly touching on what we consider to be a pleasant turn of events. The Zhang-Tapia-Dennis theory for fast convergence of the iteration sequence to zero in the Kojima-Mizuno-Yoshise algorithm was viewed by many as having the restrictive assumption of iteration sequence convergence. On the other hand, a selling point of the Ye-Güler-Tapia-Zhang theory of fast convergence of the duality gap to zero in the Mizuno-Todd-Ye algorithm was that it did not make the assumption of iteration sequence convergence. The present theory shows that the assumption of iteration sequence convergence is not restrictive in the Zhang-Tapia-Dennis theory, and the recent Gonzaga-

Tapia theory shows that the Ye-Güler-Tapia-Zhang theory is not more general than that which would result from adding the assumption of iteration sequence convergence. It seems then that fast convergence of the duality gap sequence to zero and convergence of the iteration sequence go hand in hand.

Acknowledgment

The authors thank Jorge Moré for referring us to the Ostrowski result. They also thank two anonymous referees and the co-editor for comments that led to improvements in the presentation.

References

- [1] C. G. Broyden, and J. E. Dennis Jr. and J. J. Moré. On the local and superlinear convergence of Quasi-Newton methods. *J. Inst. Math. Appl.*, 12:223-245, 1973.
- [2] A. El-Bakry, R. A. Tapia, and Yin Zhang. On the use of indicators in identifying zero variables for interior point methods. Technical Report TR91-15, Dept. Mathematical Sciences, Rice University, 1991. To appear in *SIAM Review*.
- [3] B. R. Gelbaum and John M. H. Olmsted. *Counterexamples in Analysis*. Holden-Day Inc., San Francisco, 1964.
- [4] C. Gonzaga and R. A. Tapia. On the convergence of the Mizuno-Todd-Ye algorithm to the analytic center of the solution set. Technical Report TR92-41, Dept. of Computational and Applied Mathematics, Rice University, 1992. Submitted for publication.
- [5] O. Güler and Y. Ye. Convergence behavior of some interior-point algorithms. Working Paper Series No. 91-4, Dept. of Management Sciences, The University of Iowa, 1991.
- [6] F. Hausdorff. *Mengenlehre*. Dover Publications, New York, 1944.

- [7] A. J. Hoffman. On approximate solutions of systems of linear inequalities. *Journal of Research of the National Bureau of Standards*, 49:263-265, 1952.
- [8] J. Ji, F. Potra, R. A. Tapia, and Y. Zhang. An interior-point method for linear complementarity problems with polynomial complexity and superlinear convergence. Technical Report TR91-23, Dept. Mathematical Sciences, Rice University, 1991. Submitted for publication.
- [9] M. Kojima, Y. Kurita, and S. Mizuno. Large-step interior point algorithms for linear complementarity problems. Research Report on Information Sciences B-243, Dept. of Information Sciences, Tokyo Institute of Technology, 1991.
- [10] M. Kojima, S. Mizuno, and A. Yoshise. A primal-dual interior point method for linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 29–47. Springer-Verlag, New York, 1989.
- [11] I.J. Lustig, R.E. Marsten, and D.F. Shanno. Computational experience with a primal-dual interior point method for linear programming. Technical Report SOR 89-17, Dept. Civil Eng. and O.R., Princeton University, 1989.
- [12] K. McShane. A superlinearly convergent $O(\sqrt{n}L)$ -iteration primal-dual linear programming algorithm. Manuscript, 2534 Villanova Drive, Vienna, Virginia, 1991.
- [13] K.A. McShane, C.L. Monma, and D.F. Shanno. An implementation of a primal-dual interior point method for linear programming. *ORSA J. Computing*, 1:70–83, 1989.
- [14] N. Megiddo. Pathways to the optimal set in linear programming. In Nimrod Megiddo, editor, *Progress in Mathematical programming, interior-point and related methods*, pages 131–158. Springer-Verlag, New York, 1989.
- [15] S. Mehrotra. Quadratic convergence in a primal-dual method. Technical Report 91-15, Department of Industrial Engineering and Management Science, Northwestern University, 1991.

- [16] R.D.C. Monteiro and I. Adler. Interior path-following primal-dual algorithms. Part I: linear programming. *Math. Prog.*, 44:27–41, 1989.
- [17] A. M. Ostrowski. *Solution of Equations and Systems of Equations*. Academic Press, New York, 1966.
- [18] M.J. Todd and Y. Ye. A centered projective algorithm for linear programming. *Math. of O.R.*, 15:508–529, 1990.
- [19] Y. Ye, O. Güler, R. A. Tapia, and Y. Zhang. A quadratically convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming. Technical Report TR91-26, Dept. Mathematical Sciences, Rice University, 1991.
- [20] Y. Ye, R. A. Tapia, and Y. Zhang. A superlinearly convergent $O(\sqrt{n}L)$ -iteration algorithm for linear programming. Technical Report TR91-22, Dept. Mathematical Sciences, Rice University, 1991.
- [21] Y. Zhang and R. A. Tapia. A superlinearly convergent polynomial primal-dual interior-point algorithm for linear programming. Technical Report No. 90-40, Dept. of Mathematical Sciences, Rice University, 1990. To appear in *SIAM Journal on Optimization*.
- [22] Y. Zhang and R. A. Tapia. On the convergence of interior-point methods to the center of the solution set in linear programming. Technical Report TR91-30, Dept. of Computational and Applied Mathematics, Rice University, 1991. Submitted for publication.
- [23] Y. Zhang and R. A. Tapia. Superlinear and quadratic convergence of primal-dual interior-point algorithms for linear programming revisited. Technical Report TR91-27, Dept. of Computational and Applied Mathematics, Rice University, 1991. Has also appeared in *JOTA* 73(1992), pp. 229–242.
- [24] Y. Zhang, R. A. Tapia and J. E. Dennis. On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms. Technical Report TR90-6,

Dept. Mathematical Sciences, Rice University, 1990. Has also appeared in *SIAM Journal on Optimization* 2(1992), pp. 304–324.

- [25] Y. Zhang, R. A. Tapia and F. Potra. On the superlinear convergence of interior point algorithms for a general class of problems. Technical Report TR90-9, Dept. Mathematical Sciences, Rice University, 1990. To appear in *SIAM Journal on Optimization*.